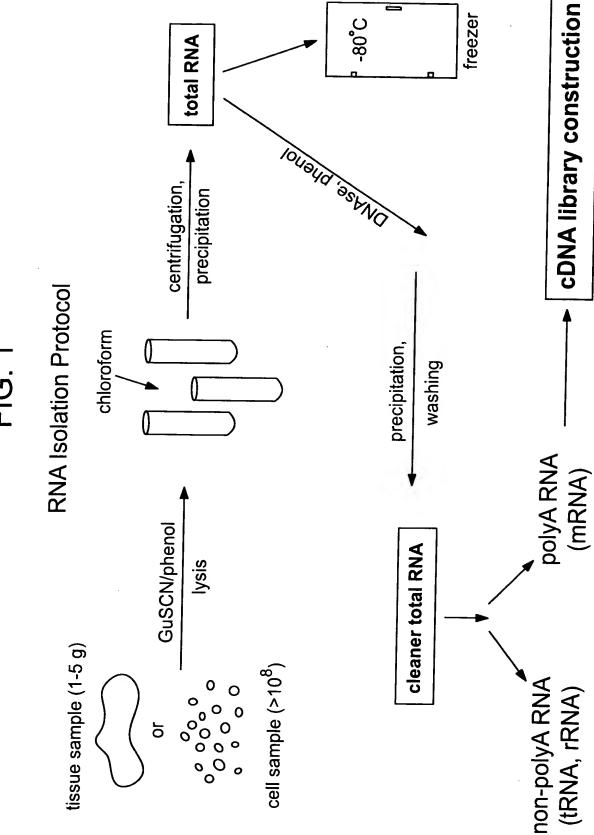
Title: Database for Storage and Analysis: Christian Central Sequences Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No. 06514069CON Applicant Lincoln et al.

FIG. 1

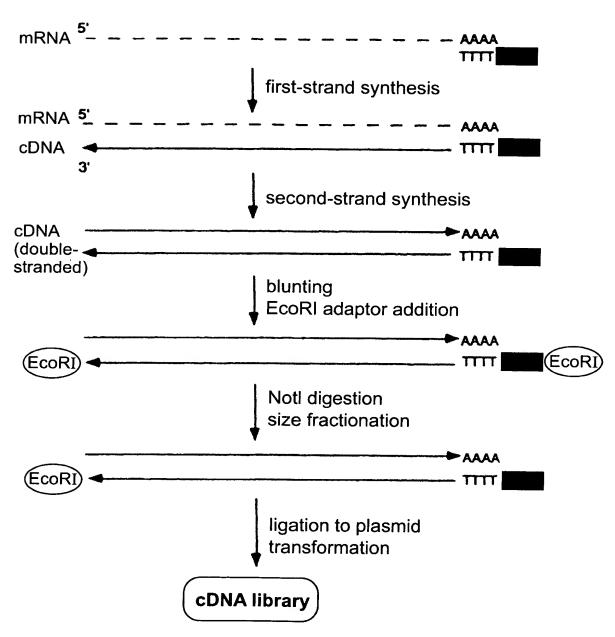


freezer

-80္နင

Analysis of Full-Length Sequences Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 2 cDNA Library Construction



Applicant Lincoln et al.

Title: Database for Storage and
Title: Database for Storage and
Title: Database for Storage and
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

Block 1 Sequence Editing Screens

Target Sequence Feature	Editing Method	Result
5' and 3' Vector	Dynamic Programming	Clip
PolyA Tail	Regular Expression	Clip
Sequencing Artifacts	Nearest Neighbor	Remove
Low Information	BLAST (S≥90)	Mask
Contamination	BLAST (S ≥ 90)	Remove
Repetitive Elements	BLAST (S≥90)	Mask
Mitochondrial	BLAST (S ≥ 90)	Remove
Ribosomal RNA	BLAST (S ≥ 90)	Remove

Applicant Lincoln et al.

Title: Database for Storage and
Title: Database for Storage and
Title: Database for Storage and
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 4

		26	56	56	56	224
ס	 -	14	14	14	14	56
Expected	IJ	14	14	14	14	56
Exp	ပ	14	14	14	14	26
	4	14	14	14	14	26
		4	ပ	G	H	
						·
		56	56	56	56	224
••	-	ω	œ	8	32	26
berrant	IJ	8	8	32	8	56
Abe	ပ	8	32	8	8	56
	4	32	8	8	8	26
	-		ပ	ŋ	_	

e.g. CCCCGGGGTTTTCCCCCAAAAGGGG...

Analysis of Full-Leng quence Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 5

Rigorous Statistics:

$$E = KNe^{-\lambda S}$$
 where $S = \sum_{HSP} s(a_i, b_j)$

$$s(a,a) = 5$$
 and $s(a,b) = -4$ for $a \neq b$

FIG. 6

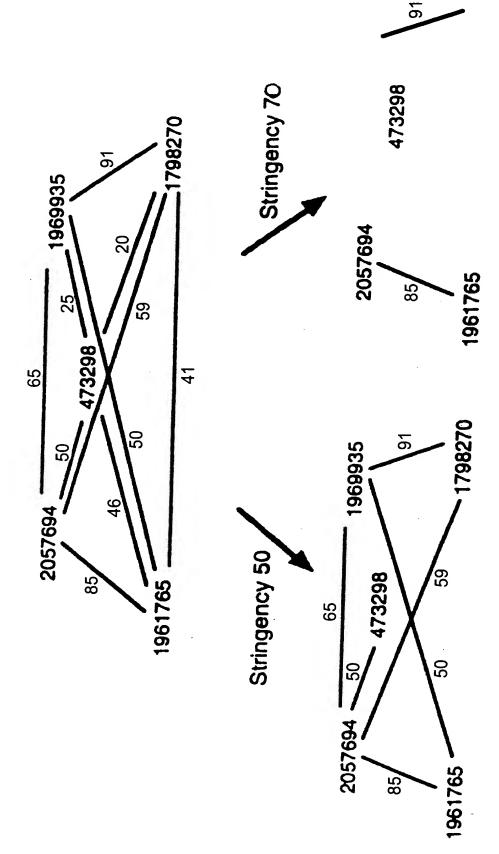
Query	Match	BLAST Score	% ID	Length	Product Score
1969935 (238)	1798270	1103	99	259	91
	2057694	765	100	234	65
	1961765	600	100	245	50
	473298	318	92	261	25
473298 (261)	2057694	634	93	234	50
10200 (201)	1961765	634	93	254	46
	1969935	318	92	238	25
	1798270	254	94	259	18

Alignments

1969935	
1798270	
2057694	
1961765	
473298	

Applicant Lincoln et al.
Title: Database for Storage and
Title: Madysis christistics and
Title: Database for Storage and
Title: Database for Storage and
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 7



Applicant Lincoln et al.

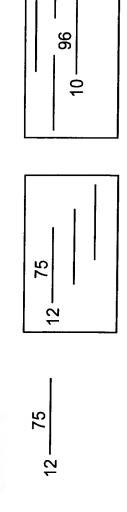
Title: Database for Storage and
II. I. I. I. Analysis of Multiple Linguis Sequences
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 8

Creating a Master Cluster

Example: Cluster - 12 (singleton), Cluster 2, and Cluster 1 all contain representative clones with PS ≥ 40 to Gene X.

Gene X



Cluster ID: 1 Rep. Clone: 10

PS = 96

Cluster ID: -12 Rep. Clone: 12

PS = 75

Rep. Clone: 7

Cluster ID: 2

PS = 100

Analysis of Fundamental Sequences Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 9

Naming a Cluster

Q: Does any member of the cluster have a match to GenBank?

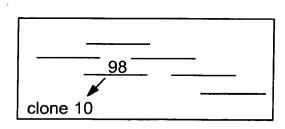
5 clones (solid line) match GI #1 1 clone (dashed) matches GI #2



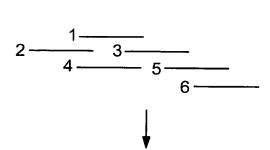
Q: Which GI is represented the most in the cluster?

A: GI #1 is represented the most.



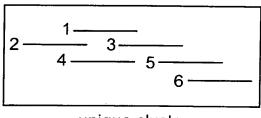


A: Clone 10 has a Product Score of 98, so the cluster is named after this representative clone.



NO

Q: Which clone has the lowest Clone ID for this cluster?



unique cluster

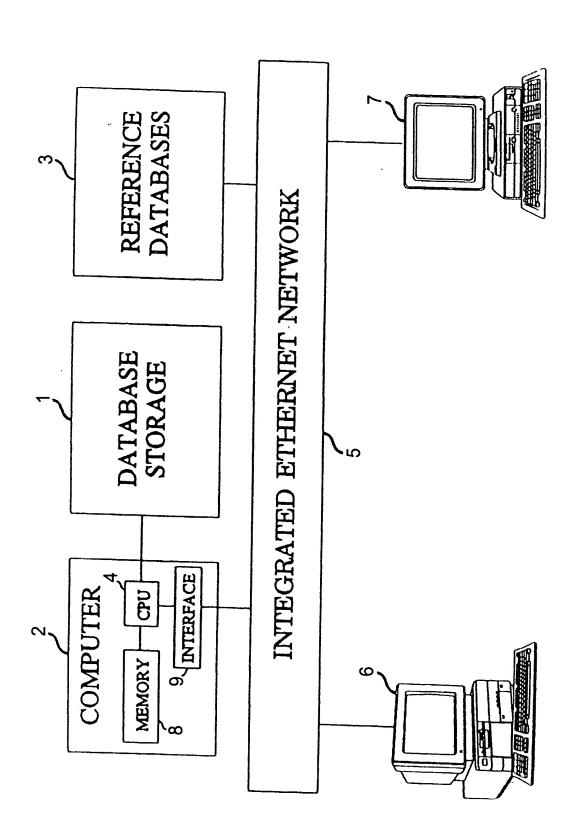
A: Clone 1 has the lowest ID, so the unique cluster is named after this representative clone.

Applicant Lincoln et al.

Title: Database for Storage and
Title: Database for Storage and
This is a finite of the concess.

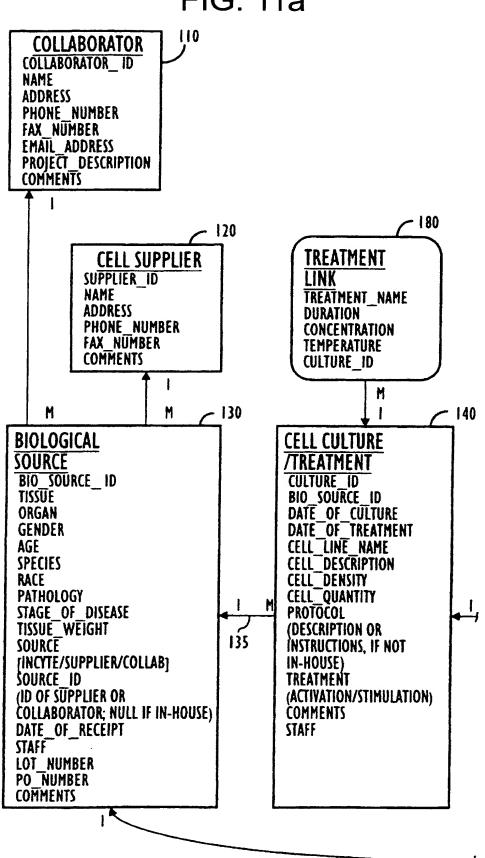
Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 10

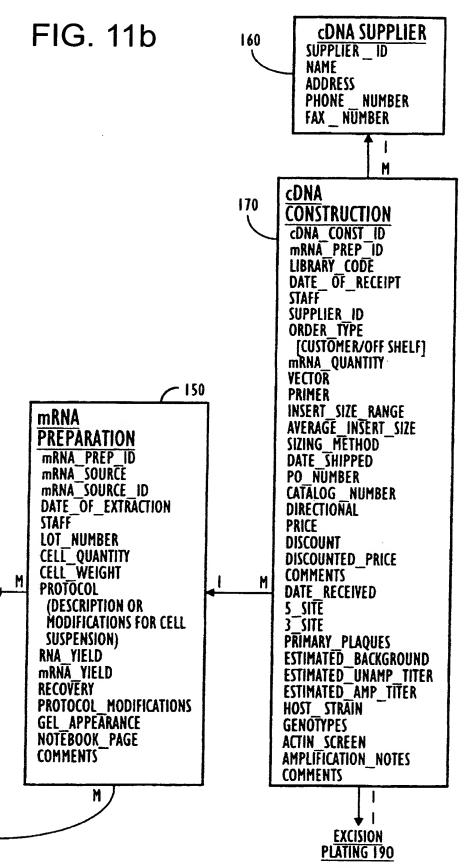


Title: Database Storage and Analysis of Flength Sequences Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 11a



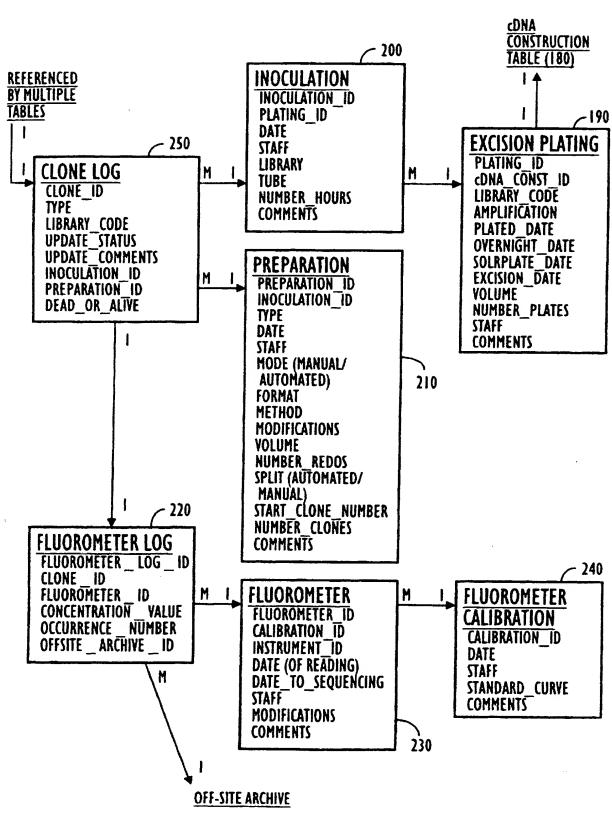
Analysis of Full-Lengt guences Filing Date: Septembe 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON



SUBSTITUTE SHEET (RULE 26)

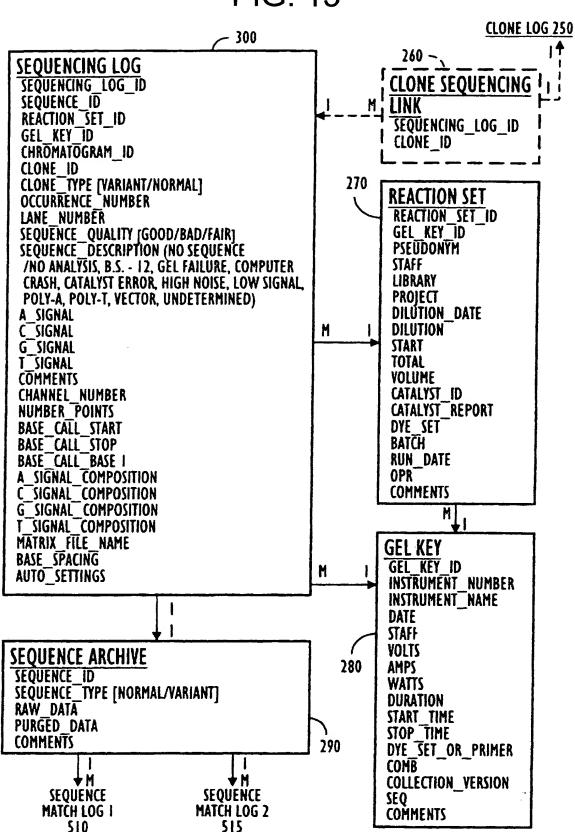
Title: Database for Storage and Analysis of Full-Such Sequences Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 12

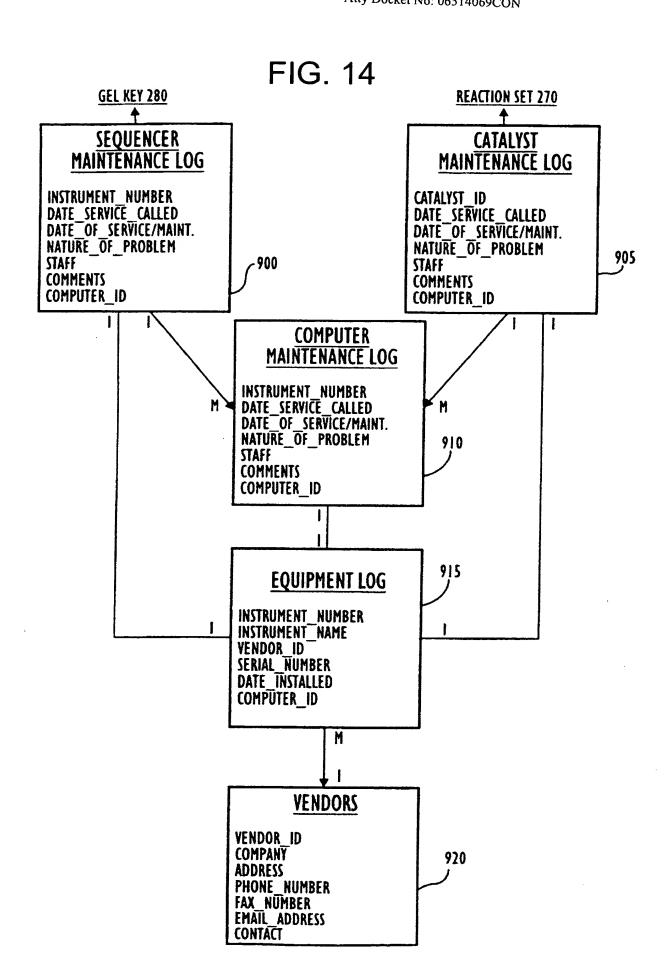


Analysis of Fundamental Sequences Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 13

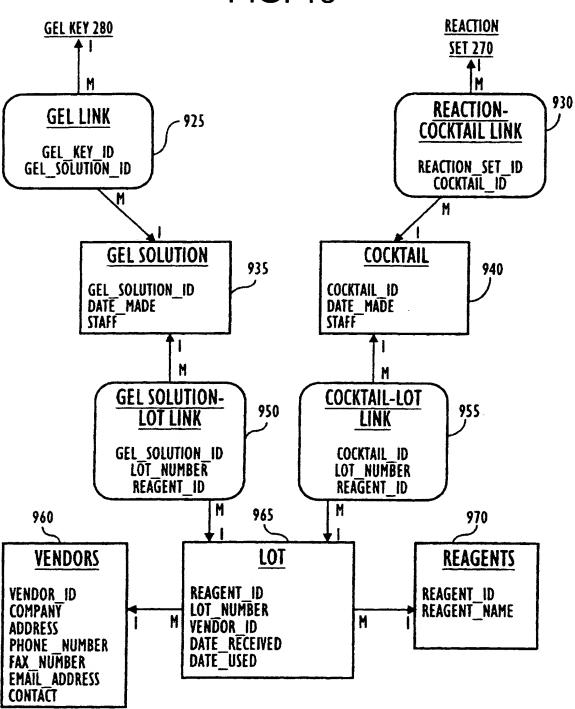


Analysis of Full-Len quences Filing Date: September 9, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON



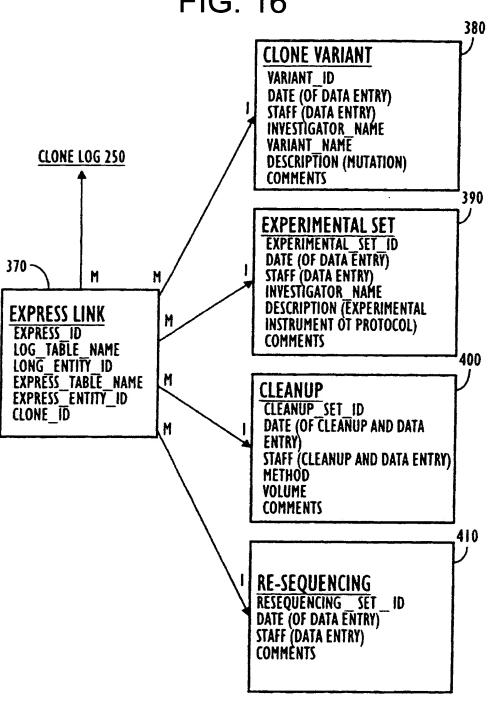
Title: Database torage and Analysis of Full gth Sequences Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 15



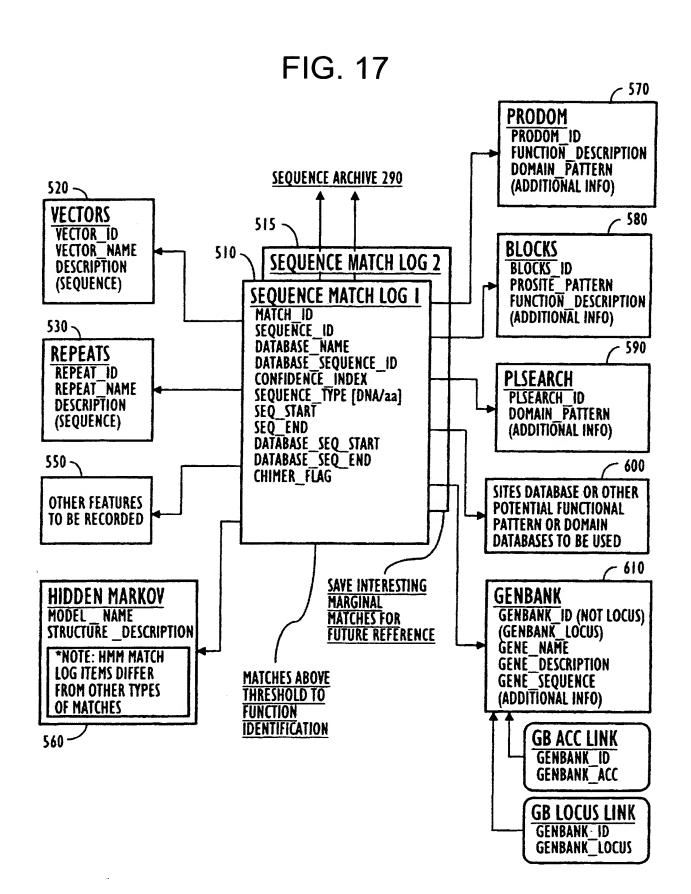
Title: Databas
Storage and
Analysis of Fundament Storage and
Analysis of Fundament Sequences
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 16



Applicant Lincoln et al.

Title: Database orage and
Analysis of Full-Length Sequences
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON



Title: Database for prage and Analysis of Full-I n Sequences Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON

FIG. 18

